

IA, responsabilité et facteurs psychologiques dans les investissements

Entretien avec Martin Nimbach

NBK Legal | Octobre 2025

I. Introduction – Contexte et parcours personnel

Question : *Cher Martin, merci beaucoup de prendre le temps de répondre à nos questions sur le thème « IA et éthique » ! Pour commencer, peux-tu te présenter brièvement et nous dire quel est ton métier ? Quel est ton lien avec l'IA et/ou les questions éthiques ?*

Martin Nimbach :

Je viens du domaine de l'ingénierie et j'ai travaillé pendant de nombreuses années à l'interface entre la technologie et l'organisation, dans des groupes, des start-ups et des équipes de recherche interdisciplinaires. Aujourd'hui, je m'intéresse principalement au lien entre l'IA, l'architecture, le leadership et les modèles neurocognitifs, c'est-à-dire les modèles qui décrivent comment la pensée, les émotions et les décisions prennent forme dans le cerveau. Ils relient la psychologie (perception, pensée, émotions, décisions) aux neurosciences (activité des cellules nerveuses, réseaux, neurotransmetteurs, etc.).

Question : *Que signifie l'IA pour toi ?*

Martin Nimbach :

Pour moi, l'IA est un nouvel outil très flexible. Elle permet de traiter des sujets complexes, même sans recourir à la programmation traditionnelle. Les grands modèles linguistiques (Large Language Models ; LLM) reflètent une grande partie de ce qui se trouve sur Internet et donnent un accès rapide à la connaissance. Cependant, leurs résultats ne sont pas infallibles – il faut donc toujours les utiliser avec prudence et les vérifier.

II. Observations et questions systémiques

Question : *Comment vis-tu le débat actuel sur l'IA – dans ton secteur, ton domaine d'expertise, ton environnement ?*

Martin Nimbach :

Dans de nombreuses organisations, il existe un champ de tension entre le battage médiatique, le surmenage et la désinformation. L'IA est souvent mise en œuvre, mais rarement vraiment *intégrée*.

Question : *Selon toi, qu'est-ce qui est négligé ou sous-estimé ?*

Martin Nimbach :

La qualité structurelle des modèles : les biais, c'est-à-dire les distorsions qui influencent les résultats fournis par l'IA, l'interprétabilité et la gouvernance. De nombreux systèmes ne sont tout simplement pas assez robustes pour être utilisés de manière productive.

Question : *Y a-t-il des questions éthiques qui te préoccupent particulièrement ?*

Martin Nimbach :

Oui, en particulier la sécurité juridique lors de l'utilisation des données qui servent de base à l'entraînement des modèles.

III. Responsabilité et conception

Question : *Qui est responsable de l'utilisation éthique de l'IA – et qui en assume réellement la responsabilité ?*

Martin Nimbach :

D'un point de vue technique, il faut des rôles clairement définis : des instances d'audit, des comités d'éthique et une architecture de sécurité robuste. Tout le reste n'est que vœu pieux.

Question : *De quoi avons-nous besoin pour rendre les principes éthiques efficaces ?*

Martin Nimbach :

Des normes contraignantes, des tests transparents et des obligations d'information dans les domaines d'application critiques.

Question : *As-tu tes propres principes ?*

Martin Nimbach :

Oui, notre cadre s'appelle « CLEAR » : Clarity – Leadership – Ethics – Awareness – Resilience (Clarté – Leadership – Éthique – Sensibilisation – Résilience).

IV. Pratique et expérience

Question : *As-tu un exemple concret de bonne ou mauvaise pratique en matière d'IA ?*

Martin Nimbach :

Dans le cadre d'un projet technologique, une IA a été utilisée pour la sélection des ressources humaines, mais personne n'a vérifié les données d'entraînement. Cela a entraîné des biais systématiques qui n'ont été remarqués que tardivement.

Question : *Qu'est-ce qu'une « bonne pratique » ?*

Martin Nimbach :

Une architecture compréhensible, des audits réguliers, l'explicabilité dès la conception (c'est-à-dire que l'IA doit pouvoir expliquer de manière compréhensible *comment* et *pourquoi* elle parvient à un certain résultat), des mécanismes de contrôle redondants (c'est-à-dire plusieurs systèmes ou procédures se sécurisant mutuellement) et une invite bien pensée. Cela semble technique, mais cela signifie en substance que l'IA doit être compréhensible et doublement sécurisée.

V. Perspectives et attitude

Question : *Qu'est-ce qui te rassure dans l'utilisation de l'IA ?*

Martin Nimbach :

L'intégration d'instances de contrôle basées sur des règles, comparables au cortex préfrontal, la partie du cerveau responsable du contrôle et de la réflexion.

Notre intuition s'apparente à un modèle linguistique à grande échelle, mais la nature nous a dotés d'un contrôle des impulsions qui agit comme un filtre sur nos actions. Un modèle de contrôle similaire, assorti de règles éthiques, devrait également être mis en œuvre dans les systèmes d'IA avant que ceux-ci ne produisent des résultats.

Question : *Qu'est-ce qui te préoccupe ?*

Martin Nimbach :

Une utilisation naïve et non critique des LLM.

Question : *Que doit-il se passer au cours des cinq prochaines années ?*

Martin Nimbach :

Nous avons besoin de normes de sécurité contraignantes, de principes d'architecture éthiques et de mécanismes d'audit interdisciplinaires.

Que signifie l'engouement pour l'IA pour les investisseurs ? – 5 enseignements approfondis pour les investisseurs avisés

1. La confiance n'est pas une caractéristique de l'IA, mais une compétence managériale. La confiance dans l'IA ne naît pas de ses performances techniques, mais d'un management responsable. Aucun modèle ne peut inspirer confiance si l'équipe n'est pas capable de prendre des décisions de manière transparente et responsable. Tout dépend donc de la manière dont *les gens* travaillent avec l'IA.

Les investisseurs devraient se poser les questions suivantes :

- Qui assume la responsabilité finale des décisions prises par le modèle ?
- Existe-t-il des procès-verbaux documentés pour les déclarations critiques de l'IA ?
- Y a-t-il des désaccords actifs au sein de l'équipe et comment sont-ils gérés ?

2. Un bon modèle sans audit est un atout incalculable.

Un modèle fonctionnel ne suffit pas à lui seul : il doit être explicable, testable et auditable. Sans analyses de biais, niveaux de vérification et processus de contrôle externes, il en résulte un manque de transparence structurel.

Les investisseurs devraient se poser les questions suivantes :

- Comment les biais sont-ils détectés, documentés et réduits ?
- Qui effectue les audits – en interne, en externe, régulièrement ?
- Existe-t-il un manuel d'intervention en cas d'incident pour les défaillances du modèle ?

3. L'évolutivité commence par la gouvernance, et non par la puissance des

GPU. De nombreuses start-ups spécialisées dans l'IA investissent dans la formation de modèles, mais négligent l'organisation qui doit les soutenir. Une IA évolutive nécessite une gouvernance évolutive, avec des rôles clairs, une gestion des versions et une architecture éthique.

Les investisseurs devraient se poser les questions suivantes :

- Quels sont les rôles existants en matière de gouvernance et d'éthique de l'IA ?
- Existe-t-il des protocoles pour le suivi des versions des modèles et des responsabilités ?
- L'entreprise est-elle conforme à la réglementation (par exemple, en ce qui concerne la loi sur l'IA) ?

4. L'expertise n'a de valeur que si elle se remet en question.

De nombreuses équipes techniques sont excellentes, mais cognitivement homogènes. L'interdisciplinarité est un facteur de sécurité. Seules des équipes diversifiées bénéficiant d'une sécurité psychologique peuvent identifier les faiblesses à un stade précoce et en tirer des enseignements.

Les investisseurs devraient se poser les questions suivantes :

- Dans quelle mesure l'équipe est-elle diversifiée en termes de culture professionnelle, de perspectives et d'expérience ?
- Quand des critiques externes ont-elles été prises en compte pour la dernière fois, et comment y a-t-on réagi ?
- Existe-t-il des formats internes pour des analyses structurées des erreurs ?

5. L'IA n'est pas un avantage intrinsèque, mais un amplificateur.

L'IA n'améliore rien en soi, elle renforce ce qui est déjà présent dans le système. Les bonnes organisations deviennent plus efficaces, les organisations dysfonctionnelles plus dangereuses.

Les investisseurs devraient se poser la question suivante :

- Quelles faiblesses ou quels conflits pourraient être exacerbés par l'IA ?
- La technologie sert-elle à clarifier les choses ou seulement à accroître l'efficacité ?
- Qu'a fait l'entreprise pour adapter sa structure, son leadership et sa culture à l'IA ?

Ces cinq perspectives offrent aux investisseurs une base d'évaluation globale : techniquement fiable, psychologiquement fondée et structurellement bien pensée. Elles aident à comprendre non seulement le produit, mais aussi l'organisation dans son ensemble, et à distinguer la véritable résilience de la simple performance.

Merci beaucoup pour ton temps, cher Martin !